Clinical Data Exchange Standards and Vocabularies for Messages

Stanley M. Huff, MD, Intermountain Health Care, Salt Lake City, UT

Motivation for the creation of electronic data interchange (message) standards is discussed. The ISO Open Systems Interface model is described. Clinical information models, message syntax and structure, and the need for a standardized coded vocabulary are explained. The HIPAA legislation and subsequent HHS transaction recommendations are reviewed. The history and mission statements of six of the most popular message development organizations (MDOs) are summarized, and the data exchange standards developed by these organizations are listed. The organizations described include Health Level Seven (HL7), American Standards for Testing and Materials (ASTM) E31, Digital Image Communication in Medicine (DICOM), European Committee for Standardization (Comité Européen de Normalisation), Technical Committee for Health Informatics (CEN/TC 251), the National Council for Prescription Drug Programs (NCPDP), and Accredited Standards Committee X12 Insurance Subcommittee (X12N). The locations of Internet web sites for the six organizations are provided as resources for further information.

INTRODUCTION

Clinical data is stored electronically in literally hundreds of different kinds of information systems. To provide optimal care for patients, the data needs to be shared between systems. Computer to computer interfaces can help to make this information available when and where it is needed. Message standards define the structure and content of data that can be exchanged between systems, as well as the policies and procedures that guide the exchange.

Donald Simborg, Clement McDonald, Ed Hammond and others have championed the idea of "best of breed" component architectures for nearly 15 years[1-5]. In this paradigm, clinical information is made available by assembling a network of integrated health care applications (components), with each component representing the best available solution for a particular information system function. Computer to computer interfaces are essential in this architecture, making it possible to share information among the networked systems.

Interfaces are also essential for communication of data between different healthcare enterprises and between private institutions and governmental agencies. Data sharing between institutions is important for public health reporting, for pooling of data for clinical research and outcomes studies, and for billing and patient account management.

In the early to mid 1980's, MDOs were formed to define standards for the electronic exchange of clinical data. The goal of the standards organizations was to reduce the cost of creating, installing, and maintaining interfaces. The MDOs have been quite successful. Probably more than 80% of newly installed medical interfaces adhere to one of the message standards. Given the prevalence of these interfaces, it seems appropriate to discuss some fundamental principles related to computer to computer messaging and to give a brief summary of six of the most popular MDOs.

FUNDAMENTALS OF MESSAGING

The goal of medical interfaces is the unambiguous transmission of information between medical systems. A basic model for all interfaces has been described by the International Standards Organization Open Systems Interconnection (ISO-OSI) specification. As shown in Figure 1, this model decomposes the process of messaging into seven logical layers or levels. Each layer

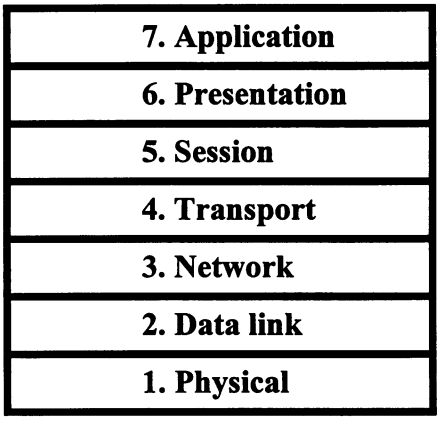

Figure 1: The ISO-OSI 7 Layer Model.

defines a specific function necessary for passing messages between systems. The definitions for the seven layers are as follows:

- Physical link – defines the electrical and mechanical aspects of interfacing to a physical medium for transmitting data. This layer includes the software device driver for each communications device plus the hardware itself – interface devices, modems, and communication lines.
- Data link – establishes an error-free communications path between network nodes over the physical channel, frames messages for transmission, checks

1091—8280/98/$5.00

integrity of received messages, manages access to and use of the channel, and ensures proper sequence of transmitted data.

- Network control - addresses messages, sets up the path between communicating nodes, routes messages across intervening nodes, and controls the flow of messages between nodes.
- Transport – provides end-to-end control of a communication session once the path has been established, allowing the processes to exchange data reliably and sequentially, independent of which systems are communicating or their location in the network.
- Session control – establishes and controls system dependent aspects of communications sessions between specific nodes in the network and bridges the gap between the services provided by the transport layer and the logical functions running under the operating system in a participating node.
- Presentation control – encoded data that has been transmitted is translated and converted into formats which enable display on terminal screens, or enable storage of the data in a database – forms that can be understood and directly manipulated by users.
- Application/User – services are provided that directly support user and application tasks and overall system management. Examples of services and applications provided at this level are resource sharing, file transfers, remote file access, database management, and network management.

The seven-layer model makes it possible to produce modular software for interfaces. As long as components at each level conform to the model, changes or advances in technology cause changes only in the layer(s) involved in that specific function. Thus, a change in network technology like moving from coaxial cable to fiber optic cable does not cause changes at the presentation or application layers.

In the early days of interfacing there were no readily available implementations of the lower layers and it was a challenge just to get a bit stream to pass between two systems. People spent time worrying about RS-232 connectors, wire sizes, asynchronous versus synchronous modems, etc. Nowadays, there are reliable, standardized, economical implementations of levels 1-5 such as Ethernet, TCP/IP, IIOP, and Berkley socket connections, and attention is focused almost entirely on levels 6 and 7 of the OSI model.

There are now also good implementations for the level 6 encoding-decoding of message content. These include Abstract Syntax Notation 1 (ASN.1[6, 7]), Standard Generalized Markup Language (SGML[8]) and its subtypes HyperText Markup Language (HTML[9]) and Extensible Markup Language (XML[10]), and Interface Definition Language (IDL). However, these tools were not available at the time the current set of medical messaging standards were created, so the existing medical standards have defined their own message syntax. Common formats in use today include fixed length messages, delimited messages, and name-value pair (tagged) messages. As discussed briefly below, an active area of investigation for many standards organizations is how they can take better advantage of standards like SGML/XML, ASN.1, or IDL.

Given that implementations of the lower levels of the OSI model are becoming more uniform and reliable, the major work of defining medical interfaces is focused on level 7. This involves the following activities:

- Defining the business needs and circumstances that are the basis for data exchange between medical systems. "Use cases" can help capture this kind of information. For example, there may be a need to share orders between a hospital information system and a clinical laboratory system. A use case would describe the clinical scenario including the actors, actions, and motivations that surround an order transaction.
- Describing the specific real world events that "trigger" the exchange of messages. For example, entry of a result into a laboratory system may be the real world event that triggers sending a result message to an emergency room system.
- Describing the attributes and relationships of entities (objects) that are the subjects of the communication. This shared public model of objects, attributes, and relationships is the basis for semantic understanding of message contents. The model contains descriptions of patients, orders, addresses, organizations, patient locations, facilities, observations, actions, etc.
- Specifying the exact information content for each type of message, including the fields and collections of fields (sets or segments) to be sent, their data type or format type, and their allowed values. The information content of a message should be based on previously defined business and information models. In the best designs, specifying the message content is kept separate from decisions about message syntax, which is really a layer 6 issue.
- Specifying the sequence in which related messages will be passed to accomplish a business need. For example, there is the expectation, based on the business model, that a laboratory system will receive a message to admit a patient to a specific bed before receiving a message to transfer the patient to a new bed.
- Defining protocols for application level acknowledgement of messages, and strategies for communicating application level errors. Application

level errors are those that occur because a receiving application can not use the message as intended. For example, an application error would occur if a system received an order for a patient who had not yet been registered on the system.

The existing message standards have been quite successful, and the number of installed interfaces continues to increase. This is due primarily to the fact that the standards *do* decrease the time and cost of implementing interfaces. The use of standard message contents, common data types, and reliable network services has resulted in less analysis and programming to install new interfaces. However, there is still a good deal of variability from one installation to the next due to the unique needs of any given site as well as occasional ambiguities in the standards. This results in bilateral negotiations and agreements between the sending and receiving parties each time an interface is created. Several systems integration companies have fulfilled a market need by creating systems (gateways) that automate conversions between different message formats. These systems also manage installation, startup, and shutdown of particular interface connections.

A goal of the next generation of interfaces is to achieve a much higher level of "plug and play" interoperability between interfaces that are implementing the same standard. Complete plug-and-play interoperability would mean that interfaces could be created based solely on the standard. In the ideal world, the sending and receiving systems would have a perfect understanding of the information sent via the interface, without the need for bilateral agreements. Absolute plug-and-play may never be achieved, but it is clearly possible to get much closer than we have achieved today.

There are three areas where the interoperability of interfaces could be improved. First, interfaces need to be based on formal information models. Nearly all of the MDOs are involved in developing or already have a formal model. Model development is a large task, and existing models need to be completed and refined.

Second, interoperability would be improved if standard encoding-decoding software could be used in interfaces. HL7 has a Special Interest Group (http://www.mcis.duke.edu/standards/HL7/committees/sgml/) that is investigating how SGML and XML might be used in some future version of their standard. There is a workgroup (http://www.xmledi.com/repository/) investigating how X12 messages could be sent using XML. ASTM subcommittee E31.11 is investigating how SGML/XML could be used to support portability of the electronic health record. The other MDOs are supporting similar initiatives. Using standard technology for creation and parsing of messages will further reduce variability in messages, as well as the cost of creating new interfaces.

Third, the coded terms used in messages needs to be standardized. At the current time, it is estimated that somewhere between 75% and 90% of the work and expense of implementing a computer to computer interface is consumed in aligning vocabulary between the two interacting systems. Plug-and-play interoperability means that for each coded data element in a message there must be a known set of concepts (meanings) that are the allowed or expected values for that field. Ultimately this means that every coded field in a message must have a domain specification that can be resolved to a finite set of concepts. Furthermore, it is recognized that there is a strong interdependency between the vocabulary used in messages and the information model (as expressed by message structure). Recognition of this interdependency has led to the creation of at least one vocabulary (LOINC[11-13]) specifically for use in clinical messages. For better or worse, most of the codes used in administrative transactions are mandated by governmental regulations, including the use of ICD-9 CM codes, DRG codes, and CPT codes. Other coding initiatives are underway as well. DICOM has incorporated the use of SNOMED International[14] in its messages via the SNOMED-DICOM Microglossary[15]. Both HL7 and ASTM have subcommittees that are actively working on vocabulary tasks. Some of the important issues being discussed are the business and organizational relationships that would allow proprietary vocabularies to be used in messages.

Progress in the areas of information modelling, use of standardized message encoding languages like SGML or XML, and standardization of the coded terms used in messages should lead to a new generation of interfaces that are approaching plug-and-play compatibility.

HIPAA LEGISLATION

The United States Health Insurance Portability and Accountability Act of 1996 (HIPAA) included provisions that require the Secretary of Health and Human Services (HHS) to adopt standards for the electronic exchange of administrative and financial health care transactions. The legislation also mandates the adoption of standards for exchange of clinical data, but the deadline for doing so is later than for administrative and financial transactions. Ultimately, all parties exchanging data of the type covered by the legislation will be required to use the standards adopted by HHS. The HHS is proposing the following standards for adoption:

1. Health care claim and equivalent encounter:
   - Retail drug: NCPDP Telecommunication Claim version 3.2 or equivalent NCPDP Batch Standard Version 1.0
   - Dental claim: ASC X12N 837 - Health Care Claim: Dental

- Professional claim: ASC X12N 837 - Health Care Claim: Professional
- Institutional claim: ASC X12N 837 - Health Care Claim: Institutional

2. Health care payment and remittance advice: ASC X12N 835 - Health Care Payment/Advice
3. Coordination of benefits:
   - Retail drug: NCPDP Telecommunication Standard Format version 3.2 or equivalent NCPDP Batch Standard Version 1.0
   - Dental claim: ASC X12N 837 - Health Care Claim: Dental
   - Professional claim: ASC X12N 837 - Health Care Claim: Professional
   - Institutional claim: ASC X12N 837 - Health Care Claim: Institutional
4. Health claim status: ASC X12N 276/277 - Health Care Claim Status Request and Response
5. Enrollment and disenrollment in a health plan: ASC X12 834 - Benefit Enrollment and Maintenance
6. Eligibility for a health plan: ASC X12N 270/271 - Health Care Eligibility Benefit Inquiry and Response
7. Health plan premium payments: ASC X12 820 - Payment Order/Remittance Advice
8. Referral certification and authorization: ASC X12N 278 - Health Care Services Review - Request for Review and Response

HHS is in the process of forming a proposal for a claims attachment standard also. The attachment standards are likely to be drafted so that health care providers using HL7 for their in-house clinical systems would be able to send HL7 clinical data to health plans in association with one or more ASC X12N transactions.

As can be seen, all but one of the proposed transactions are from ASC X12 organization. The remaining transaction, for retail drug claims, is from NCPDP. For several U.S. based standards organizations the next year or so will be critical as HHS determines which standard(s) will be adopted for clinical data exchange.

SPECIFIC MDOS

The remainder of this article gives a brief history for six of the most popular medical MDOs, a list of the messages they support, and the addresses of their web sites. For the sake of brevity, I have not included descriptions of any of the instrument interface standards or of the medical information bus standards. The majority of the information on the MDOs was obtained from web sites, from information pamphlets, or by participation of the author in the organization. For information about current activities and new initiatives the reader is referred to the web site of the organization of interest. The organizations are listed in alphabetical order and include ASTM, CEN/TC 251, DICOM, HL7, NCPDP, and X12N.

ASTM

History: ASTM was established as a not-for-profit standards organization in 1898. ASTM is accredited by the American National Standards Institute (ANSI), and has 132 standards writing committees. ASTM's standards development activities cover such diverse areas as metals, paints, plastics, and consumer products. Committee E-31 on Healthcare Informatics was established in 1970. The E 1238 standard was published in 1988.

Mission/Scope: Committee E-31 "develops standards for health information systems designed to assist vendors, users and anyone interested in systematized health information. The current standards address architecture, content, portability, format, privacy, security and communications."

Message Standards: E1238 Specification for Transferring Clinical Observations between Independent Computer Systems.

Comment: ASTM standard E1238 was the first published consensus standard for the transfer of clinical data between independent computers. It is currently most often used for batch and reference laboratory interfaces. E1238 is technically aligned with Chapter 4 - Observations and Results of HL7.

Web site: http://www.astm.org/COMMIT/e-31.htm

CEN/TC 251

History: The parent organization of TC 251 is CEN. CEN is the European analogue of ANSI, but with some important differences. CEN develops standards while ANSI only coordinates standards. CEN often funds the development of standards, while vendors and interested users and participants who fund themselves develop standards in the U.S. As described on its web page "CEN's aims, expressed in its Statutes, are to draw up voluntary European Standards and promote corresponding conformity of products and services in areas other than electrotechnical and telecommunications."

Mission/Scope: The mission of CEN/TC 251 is "to develop standards that enable compatibility and interoperability between independent systems in healthcare." In addition, Working Group IV - Technology for Interoperability has as its scope "to develop and promote standards that enable the interoperability of devices and information systems in health informatics. The scope covers 3 main areas: 1) intercommunication of data between devices and information systems, 2) integration of data for multimedia representation, and 3) communication of such data between source departments and other legitimate users elsewhere in the healthcare sector, in order to facilitate electronic healthcare record provision."

Message Standards: ENV 1613 Messages for exchange of laboratory information, ENV 12052 Medical Imaging Communication, ENV 12538 Messages for patient referral and discharge, ENV 12539 Request and report messages for diagnostic service departments, ENV 12612 Messages for the exchange of healthcare administrative information.

Comment: U.S. residents can purchase CEN documents through ANSI.

Web site: http://www.centc251.org/

DICOM

History: In 1983, the American College of Radiology (ACR) and the National Electrical Manufacturers Association (NEMA) formed a joint committee called the ACR-NEMA Digital Imaging and Communications Standards Committee. Originally only NEMA and ACR members could participate, but changes have since been made that allow all interested parties to participate. The DICOM Version 1.0 standard was published in 1985, with Version 2.0 published in 1988, and Version 3.0 in 1993.

Mission/Scope: The initial mission of the committee "was to find or develop an interface between imaging equipment and whatever the user wanted to connect. In addition to specifications for the hardware connection, the standard to be developed was to include a dictionary of the data elements needed for proper image display and interpretation." The scope of the committee has expanded greatly since that time. Current versions or supplements to the standard support networked environments, and include specifications for radiology, ultrasound, and visible light images. Numeric, textual, and coded clinical data can also be sent via DICOM.

Message Standards: Digital Imaging and Communications in Medicine (DICOM), Parts 1-9; Supplement 15: Visible Light Image for Endoscopy, Microscopy, and Photography; Supplement 23: Structured Reporting.

Comment: DICOM has been very successful. It is essentially the only standard for the transfer of medical images. Version 3 of the standard is based on an object-oriented model, and SNOMED International is used as the vocabulary for coded data.

Web site:
http://www.xray.hmc.psu.edu/dicom/dicom_home.html

HL7

History: The first meeting of HL7 was held in March 1987. Working group meetings are held at least three times each year. Version 1.0 of the standard was published in October 1987, and version 2.0 was published in 1989. HL7 became an ANSI Accredited Standards Developing Organization in June 1994. Version 2.3 was approved by ANSI as an American National Standard in May 1997.

Mission/Scope: The mission of HL7 is "to provide standards for the exchange, management and integration of data that support clinical patient care and the management, delivery and evaluation of healthcare services. Specifically, to create flexible, cost effective approaches, standards, guidelines, methodologies, and related services for interoperability between healthcare information systems."

Message Standards: The HL7 standard describes messages for: Patient Administration - Admit, Discharge, Transfer and Demographics; Orders for Clinical Observations, Pharmacy, Dietary and Supplies; Patient Accounting and Charges; Clinical Observation Report Messages; Healthcare Application Master Files; Document Management Services and Resources; Appointment Scheduling; Patient Referral; and messages to support problem-oriented records.

Comment: HL7 is the most commonly used standard in the U.S. for numeric, textual, and coded clinical data. Messages are sent in an ASCII delimited format. HHS may recommend HL7 as part of the HIPAA mandated transaction set for sending clinical data associated with claims attachments. HL7 has a number of international affiliates in countries such as Australia, Canada, England, Germany, the Netherlands and New Zealand.

Web site: http://www.hl7.org/

NCPDP

History: NCPDP is a non-profit, standards development organization based in Phoenix, AZ. Membership is open to individuals and organizations from all segments of the third-party prescription drug program industry. The first annual meeting of NCPDP was in 1978.

Mission/Scope: NCPDP "works to create and promote data interchange and processing standards to the pharmacy services sector of the health care industry and to provide a continuing source of accurate and reliable information that supports the needs of their membership."

Message Standards: Telecommunication Standard Format Version 3.2, 3.3-4.2; Compound Transaction Standard Version 1.0; Billing Unit Standard Version 1.4; Batch Standard Version 1.0; Standard Claims Billing Tape Format Version 2.0; Standard Diskette Billing Format Version 2.0; Magnetic Stripe Standard Format Version 2.0; Manufacturer Rebate Standard; Member Enrollment Standard Format Version 1.0; SCRIPT Electronic Prescription Standard; Prior Authorization Standard 1.0.

Comment: NCPDP is the dominant standard for retail drug related transactions. Telecommunication Standard Format Version 3.2 and Batch Standard Version 1.0 have been recommended by HHS as part of the HIPAA mandated transaction set.

Web site: http://www.ncpdp.org/index.htm

ASC X12N

History: The American National Standards Institute (ANSI) chartered the Accredited Standards Committee (ASC) X12 in 1979. ASC X12 develops, maintains, interprets, publishes, and promotes the proper use of American National Standards and UN/EDIFACT international standards for electronic data interchange (EDI). ASC X12N was approved as a subcommittee of ASC X12 in February 1991.

Mission/Scope: "The principal responsibilities of the X12N Insurance Subcommittee are development and maintenance of X12 standards, UN/EDIFACT Messages, standards interpretations and guidelines, as they relate to all aspects of insurance and insurance-related business processes including, but not limited to, property, casualty, health care, life, annuity, reinsurance, pensions and reporting to regulatory agencies."

Message Standards: The pertinent healthcare standards are mentioned in the comment below.

Comment: X12 and X12N are the dominant standards for electronic commerce in the U.S. ASC X12N 270/271 - Health Care Eligibility Benefit Inquiry and Response, ASC X12N 276/277 - Health Care Claim Status Request and Response, ASC X12N 278 - Health Care Services Review - Request for Review and Response, ASC X12 820 - Payment Order/Remittance Advice, ASC X12 834 - Benefit Enrollment and Maintenance, ASC X12N 835 - Health Care Payment/Advice, ASC X12N 837 - Health Care Claim have been recommended by HHS as part of the HIPAA mandated transaction set.

Web site: http://polaris.disa.org/x12/x12n/

CONCLUSION

Medical data exchange standards are now common place in the medical environment. Due to the HIPAA mandates, they will become even more common in the next few years. Choosing one or more standards for the transmission of clinical data will be an important next step in the HIPAA process. Improved information models, the use of standard message formats, and the use of standard vocabularies should lead to increased plug-and-play compatibility of interfaces in the future.

References

1. Simborg DW. Local area networks: why? What? What if? *MD Comput.* 1984;1:10-20.
2. Simborg DW. Networking and medical information systems. *J Med Syst.* 1984;8:43-7.
3. McDonald CJ, Hammond WE. Standard Formats for Electronic Transfer of Clinical Data (Editorial). *Annals of Internal Medicine.* 1989;110:333-5.
4. Hammond WE. Transferring Clinical Lab Data Between Independent Computer Systems. *ASTM Standardization News.* 1988:28-30.
5. McDonald CJ. The search for national standards for medical data exchange [editorial]. *MD Comput.* 1984;1:3-4.
6. ISO/IEC 8824. Specification of Abstract Syntax Notation One (ASN.1). . Second ed. Geneva, Switzerland: International Organization for Standardization; 1990:51.
7. ISO/IEC 8825. Specification of Basic Encoding Rules for Abstract Syntax Notation One (ASN.1). . Second ed. Geneva, Switzerland: International Organization for Standardization; 1990:17.
8. ISO/IEC 8879. Standard Generalized Markup Language. . Geneva, Switzerland: International Organization for Standardization; 1986.
9. Raggett D, Le Hors A, Jacobs I. HyperText Markup Language (HTML) 4.0 Specification. : WC3 Recommendation; 1998.
10. Bray T, Paoli J, Sperberg-McQueen CM. Extensible Markup Language (XML) 1.0. : W3C Recommendation; 1998.
11. Forrey AW, McDonald CJ, DeMoor G, et al. Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results. *Clinical Chemistry.* 1996;42:81-90.
12. Huff SM, Rocha RA, McDonald CJ, et al. Development of the LOINC (Logical Observation Identifier Names and Codes) Vocabulary. *JAMIA.* 1998;5:276-292.
13. Logical Observation Identifier Names and Codes (LOINC™) Committee. Logical Observation Identifier Names and Codes (LOINC™) Users' Guide vs. 1.0 - Release 1.0j. . Indianapolis, IN: Regenstrief Institute; 1995.
14. Côté RA, Rothwell DJ, Palotay JL, Beckett RS, Brochu L. The Systematized Nomenclature of Human and Veterinary Medicine - SNOMED International. . Nothfield, IL: College of American Pathologists; 1993.
15. Bidgood WDJ. The SNOMED DICOM Microglossary: A Controlled Terminology Resource for DICOM Coded Entry Data Elements. In: Chute CG, ed. *IMIA Working Group 6.* Jacksonville, FL; 1997.